Distinct Feature Extraction for Video-based Gait Phase Classification

[1]Minxiang Ye, [2]Cheng Yang, [1]Vladimir Stankovic, [1]Lina Stankovic, [3,4]Samuel Cheng
[1]Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK
[2]Department of Electrical Engineering and Computer Science, York University, Toronto, ON M3J 1P3, Canada
[3]Department of Electronics and Information Engineering, Tongji University, Shanghai, China
[4]School of Electrical and Computer Engineering, University of Oklahoma, Tulsa, OK 73019, USA

***Abstract*—Recent advances in image acquisition and analysis have resulted in disruptive innovation in physical rehabilitation systems facilitating cost-effective, portable, video-based gait assessment. While these inexpensive motion capture systems, suitable for home rehabilitation, do not generally provide accurate kinematics measurements on their own, image processing algorithms ensure gait analysis that is accurate enough for rehabilitation programs. This paper proposes high-accuracy classification of gait phases and muscle actions, using readings from low-cost motion capture systems. First, 12 gait parameters, drawn from the medical literature, are defined to characterize gait patterns. These proposed parameters are then used as input to our proposed *multi-channel time-series classification* and gait phase reconstruction methods. Proposed methods fully utilize temporal information of gait parameters, thus improving the final classification accuracy. The validation, conducted using 126 experiments, with 6 healthy volunteers and 9 stroke survivors with manually-labelled gait phases, achieves state-of-art classification accuracy of gait phase with lower computational complexity compared to previous solutions.[1].**

***Index Terms*—feature extraction, gait phase classification**

List of Symbols

K	Number of gait phases
P	Gait phase label
L	Length of frames in standardized gait cycle curve
λ_i	The i-th gait parameter
$V_j^z(f)$	Gait parameter value at frame f in j-th gait cycle curve for gait parameter z
$S_j^z(f)$	Gait parameter value at frame f in j-th standardized gait cycle curve for gait parameter z
$\phi(P_a, P_b)$	Distance between two gait phase labels
G^z	The gait cycle curve cluster for gait parameter z
$\mathring{S}_m^z(f)$	Gait parameter value at frame f in the m-th gait cycle curve of cluster G^z for gait parameter z
$\bar{S}^z(f)$	Gait parameter value at frame f in the averaged curve $\bar{S}$ of cluster G^z for gait parameter z
$\bar{\omega}_m^z(x, y)$	Shared weight for frames x and y in m-th gait cycle curve of cluster G^z for gait parameter z
$P_{\bar{S}^z(f)}$	Gait phase label at frame f in the averaged curve $\bar{S}^z(f)$ of cluster G^z for gait parameter z
$\mathring{S}$	Gait cycle curve
$P_{\mathring{S}_m^z(f)}$	Gait phase label at frame f in the m-th gait cycle curve $\mathring{S}$ of cluster G^z for gait parameter z
τ	Continuous edge ratio of adjacent frames
(u, v)	Feature pair that links two frames offsets by $l + uL$ and $l + vL$ from centroid frame f within an L-length sliding window
$\Re_l(u, v)$	The feature value linked by (u, v)
ϖ_f	Length of frames in sliding window centered at frame f
$Q_\varepsilon^z(P_a, P_b)$	Quality of ε-th feature candidate pair for gait parameter z to classify gait phase labels P_a, P_b
Ω	Total desired number of feature pairs for all gait parameters
$\mathcal{H}$	Combination of any two gait phase labels
r_z	Neighbor radius of two gait pattern curves for gait parameter z
ρ_f	Class probability vector at frame f
$\eta_f(a, b)$	Correlation coefficient of adjacent gait phase labels P_a, P_b at frame f
$\ddot{T}_p$	Median duration of gait phase in the training set

I. Introduction

This paper tackles the challenges of tele-rehabilitation systems, in particular gait assessment, enabled by processing of measurements from sensor-driven motion capture systems. The choice of tele-rehabilitation system should not be cumbersome for the patients. The system should facilitate quick and painless data collection while the patients perform a few standard exercises, and should limit infrastructure change especially when targeting home-based or local clinic-based rehabilitation [2], [3]. Furthermore, gait analysis algorithms for physical rehabilitation programs (see [4] and references therein) should be very accurate, irrespective of sensors used.

Sensor-driven rehabilitation systems include EMG/EEG/ECG sensor-based systems, which acquire electrical activity produced by muscles, brain dynamic state changes or clinical signals with additional respiration analysis to assess gait functionalities across many applications [5]–[9]. Other systems, such as [10]–[13], employ inertial sensors and force plates to track body movements and perform gait analysis. Alternatively, optical, video-based motion capture systems usually adopt single or multiple high-end RGB/infrared cameras to track passive (*e.g.*, retro-reflective), active (*e.g.*, LED) or high-contrast markers in order to assess gait [2], [14]–[18]. Such systems have recently

[1]This work was presented in part at *IEEE ICME-2017* [1].

become increasingly popular as decision support tools in many rehabilitation applications. However, commercial multi-camera motion analysis systems, such as VICON [14] and Qualisys [15], are limited by high cost, ample space requirements, a complicated installation process and limited portability.

With advances in optical sensing technology, relatively inexpensive and portable optical motion capture systems have gained attention [5]. For example, Microsoft (MS) Kinect [19] is proposed for motion capture for: (i) Frailty syndrome detection [20], to assess a person's mobility via Timed Up and Go (TUG) tests [21]; (ii) Gait recognition [22], to classify an entire gait sequence into classes (each class corresponds to one person), integrating color, depth and inertial sensing; (iii) Sleep apnoea detection [23], to classify types of respiratory events by tracking the patient's chest and abdominal movements; (iv) Heart rate and rhythm estimation [24], to collect biometrics under high acquisition noise; (v) Treadmill-based gait training to track foot movement of patients with neurological impairments [25]; (vi) Postural control assessment to measure compensatory step length in the mediolateral direction [26] and standing balance with good concurrent validity [27]; (vii) 3D multi-view motion tracking and model-based rendering to facilitate reliable motion analysis [28]; and (viii) Post-stroke rehabilitation assessment [4] to capture kinematics and evaluate rehabilitation progress on the 4-meter walking test [29].

Regardless of the motion capture system used, both the *quality of acquired representative joint movements of interest* and *effectiveness of distinct feature selection* for specific pathological classification, have a significant impact on gait assessment methods [30]. This is especially critical for systems relying on inexpensive motion capture systems, which need to demonstrate a comparable level of accuracy to clinical systems.

Lie group features, investigated in many action recognition applications [31], [32], have proved to be a strong feature for classifying actions. However, only long-time actions are investigated in the literature (see the review paper [33] and references therein), and the frame-wise solutions often suffer from temporal misalignment of features. Population Hidden Markov Model (pHMM) is proposed in [34] to obtain a fixed length of silhouettes within a gait cycle for aligning temporal features. However, the algorithm requires high-quality silhouette extraction, reconstruction, and additional training.

Motivated by this gap in the literature, preliminary studies, and the emergence of low-cost motion capture systems suitable for home rehabilitation, in this paper, we propose a methodology for extracting features for precise alignment of gait patterns. Note that unlike gait/action recognition [22], [31]–[33], whose task is to classify the entire gait sequence into different actions (*e.g.*, sitting, standing, walking), we are focused on classifying each frame of a gait sequence into different gait phases (classes) to accurately characterize gait motion. This in turn facilitates the diagnostics of individual gait phase for physical therapy, thus improving the quality of the overall gait assessment. Therefore, making reference in italics to Fig. 3 of the overall systems architecture, the main contributions of this paper are:

1) A video-based *gait pattern extraction methodology* designed specifically to avoid over-fitting the classifier. The proposed methodology characterizes well subject-sensitive gait patterns and reduces the amount of training samples without sacrificing the classification performance. In addition to kinematics (including gait trajectories) as is usual practice, we exploit gait phase information.
2) A globally optimal *gait phase feature extraction* method to address the problem of temporal feature misalignment across frames. The proposed method achieves a global optimal set of resulting feature pairs, which are inputs to the classifier. Furthermore, to reduce computational complexity of the above filtering-based feature extraction method, without sacrificing the classification performance, we propose an optimized feature extraction method that is computationally efficient.
3) A *gait phase reconstruction* method to avoid misclassification of the periodic time sequential gait phase labels and recover the predictions.

Building on the conference version [1], where preliminary results of the feature extraction method (point 2 above) were presented, this paper includes the following additional contributions:

1) Feature candidate pair detection method (Alg. 2)
2) Feature pair mining method (Alg. 3)
3) Demonstrating the added value of using extracted gait parameters with respect to using the more traditional approach of using normalized joint trajectories directly
4) In-depth comparison and discussion between proposed and various benchmarked schemes.

In Section II, we review related work on extracting gait phase information for gait assessment. In Section III-B1, we first briefly describe the gait phase classification problem and give an overview of the proposed system. Next, we present our proposed gait pattern extraction, gait phase feature extraction and gait phase reconstruction processes. Furthermore, we propose two mining algorithms to extract feature pairs and discuss the performance of both approaches. Finally, in Section IV, we present the experimental results in terms of classification performance and computational complexity.

II. Related Work

High-level features, such as step through length, step height speed, and step interval, are extracted in [35] via an MS Kinect, during 360°-turn analysis. A simple statistical feature thresholding method further performs the classification of normal/abnormal gait. However, no numerical classification results are presented and lack of abnormal data is a limitation of this preliminary study to determe empirical thresholds. A point-of-care gait assessment framework in [36] quantifies several gait indices and evaluates limb impairment for patients with multiple sclerosis, involving Dynamic Time Warping (DTW), Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). However, [36] reports that the

inaccuracy of the captured joint trajectories limits the validation of this preliminary study. Furthermore, no benchmarking is included. In [37], representative kinematic parameters are extracted from a unified representation via a generic full-body kinematic model to segment motion sequences into repetitive action sequences, based on the zero-velocity crossing of the feature selection. The proposed unsupervised temporal segmentation method in [37] requires manual parameter tuning of the involved unscented Kalman filter, frequency analysis and adaptive k-means clustering to achieve good performance. However, such unsupervised segmentation requires prior knowledge of the number and order of actions occurred in the motion sequences. Moreover, the method fails when a full sequence is not measured correctly. A similar video-based gait phase classification system in [2] first acquires joint trajectories by tracking 2D bull-eye paper markers using a single high-speed expensive RGB camera. Heuristic thresholding criteria performs further automatic gait phase labelling once six gait events are extracted. Motivated by [2], a 3D motion capture system is proposed in [4] to facilitate the kinematics representation by tracking 3D trajectories of retro-reflective ball markers together with image processing algorithms. Based on the high precision joint trajectories, a heuristic method [4] is proposed to extract stance and swing phases. Both [2] and [4] require adaptive parameter tuning to achieve acceptable classification accuracy.

In this paper we provide the following improvements on the prior state-of-the-art work reviewed above: (1) a comprehensive set of gait parameters, based on 3D joint trajectories, to generalize high-level kinematics across all gait phases, and (2) gait phase classification of each frame, given a gait sequence, unlike [2] which focus on detecting gait event frames, and not the gait phase periods between gait events.

In [12], a lower limb exoskeleton robot ROBIN-H1, a walking rehabilitation service for stroke patients, is used to acquire pitch orientations and angular velocities of the robot legs. The task of classifying stance and swing phases by the captured kinematics are then investigated by training a Multilayer Perceptron (MLP) neural network (NN) and an NN-based non-linear autoregressive with exogenous inputs (NARX). Results show that NARX-NN outperforms MLP-NN, but the online classification accuracy (ACC) loss is 5.7% worse than its offline version. [12] suggest that further improvement could be made by acquiring more representative walking pattern data and applying an autoencoder.

Knee joint angle and foot switches are acquired as features to detect gait events in [30]. Classification fit percentages comparison between artificial neural fuzzy inference systems (ANFIS), autoregressive models with exogenous variables (ARX), output error models (OE), NARX and other NN-based models, demonstrate that the best model is NARX with a 88.59% fit rate.

In [34], a pHMM is proposed to extract a dynamics-normalized, averaged, gait cycle by observing silhouettes of gait stances, leading to the state-of-the-art identification of the human subject. Various-length silhouettes are projected into a fixed-length sequence by training pHMM on manual silhouettes created from a subset of the Gait Challenge dataset. However, human silhouettes vary with respect to subject shape, gait speed and walking direction, and it is impractical to generate a dataset manually to cover all possible silhouettes, especially for physical rehabilitation programs. Thus, a more robust and practical high-level motion representation approach is to rely on joint trajectories.

In [38], a commercial motion capture system 'Visual3D' is employed to acquire knee angle parameters to perform four-class gait phase classification by J48 Decision Tree, Random decision forest, MLP, and Support Vector Machines (SVM) [39]. It is demonstrated that random decision forest achieves the best classification accuracy.

In Section III, we optimally extract distinct features by observing transitions of 12 gait parameters within a sliding

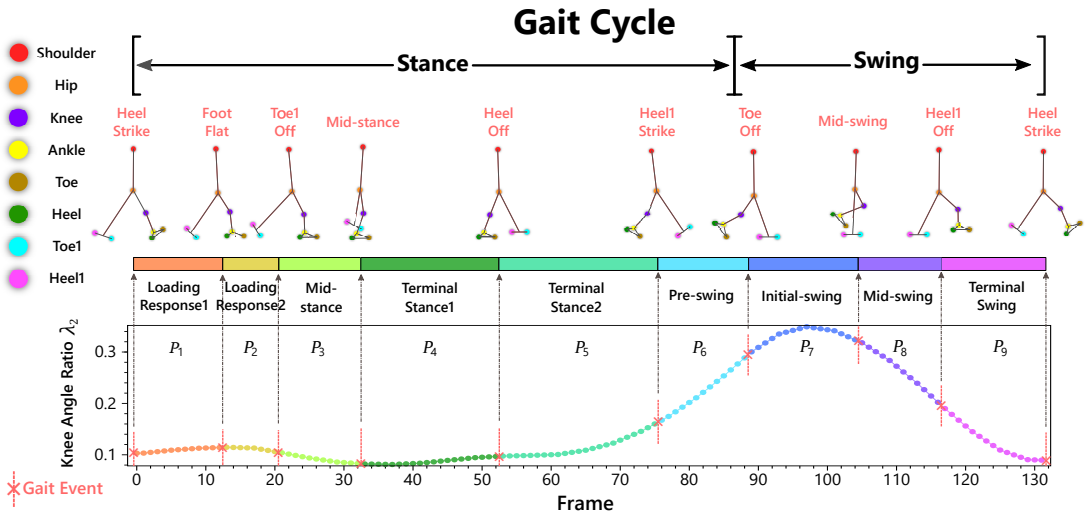

Fig. 1: A complete typical gait cycle, from a right foot heel strike until the next right foot heel strike. Circles with different colours annotate eight joints of interest listed on the left. Nine specific skeleton poses are described with their corresponding gait events, shown in red text, starting from the Heel Strike event. Colour marking of nine gait phases $P_1 - P_9$, where P_1 denotes Loading Response1, and so on. Gait phase denotes a period between two consecutive gait events. Bottom panel demonstrates a typical curve of knee angle ratio λ_2 through the whole gait cycle. Table I defines the remaining gait parameters, targeting significant gait kinematics.

window at varying gait speed and direction, unlike [12] and [30] that only rely on a single kinematics parameter at a fixed gait speed and direction. As a result, temporal feature misalignment is addressed by extracting distinct feature transitions via gait cycle standardization and clustering, mapping spatio-temporal feature transitions by random decision forest and reconstructing adjacent gait phases by fine-tuning gait events. In Section IV, we employ state-of-the-art SVM, a two-layer NN, and a NARX-NN model [12] as benchmarks for the evaluation of the classification performance.

III. Gait Phase Classification

Gait assessment reveals significant factors of abnormal gait, guiding the rehabilitation assessment and treatment, supporting clinical diagnosis and therapeutic effect evaluation. *Gait phase analysis*, as an important part of gait assessment [40], facilitates medical diagnosis [30] by analyzing movement of joints of interest.

In this paper, we propose a gait phase analysis method, which aims to accurately identify the start and end of each gait phase, that is, the time period between two consecutive gait events. This is achieved by classifying each frame of the captured video into one of $K = 9$ different gait phases (see Fig. 1), labeled as $\{P_1, \cdots, P_K\}$, in order to locate K *gait events* such as heel strike, toe off, etc.

TABLE I: Definitions & observations of the proposed 12 gait parameters. See Fig. 2 for the definition of parameters d_i and a_i.

Parameter	Value	Observation
foot distance ratio λ_1	$\frac{d_2}{d_1}$	foot switch
knee angle ratio λ_2	$\frac{a_1}{180°}$	limb support
thigh plane angle ratio λ_3	$\frac{a_2}{180°}$	femur swing
toe raise ratio λ_4	$\frac{d_3}{d_1}$	toe contact
heel raise ratio λ_5	$\frac{d_4}{d_1}$	heel contact
toe 1 raise ratio λ_6	$\frac{d_5}{d_1}$	toe 1 contact
heel 1 raise ratio λ_7	$\frac{d_6}{d_1}$	heel 1 contact
leg plane 1 angle ratio λ_8	$\frac{a_3}{180°}$	leg 1 swing
ankle angle ratio λ_9	$\frac{a_4}{180°}$	limb support
shank plane angle ratio λ_{10}	$\frac{a_5}{180°}$	tibia swing
foot angle ratio λ_{11}	$\frac{a_6}{180°}$	foot support
foot 1 angle ratio λ_{12}	$\frac{a_7}{180°}$	foot 1 support

We define 12 gait parameters $(\lambda_1, \cdots, \lambda_{12})$ and their kinematics observations, following [2] and [41], as shown in Table I. These *gait parameters* characterize gait motion during the 4-meter walking test. The change of these gait parameters during the walking test is used to perform gait phase classification. A visual representation is shown in Fig. 2. Joint trajectories [4] are used to calculate gait parameters per frame, which will, in turn, be used to perform feature extraction and classification to label the gait phases.

The block diagram of the overall proposed system is shown in Fig. 3, comprising 3 major steps: (1) gait pattern extraction, (2) gait phase feature extraction, and (3) gait phase reconstruction. These are described in the following three subsections, respectively.

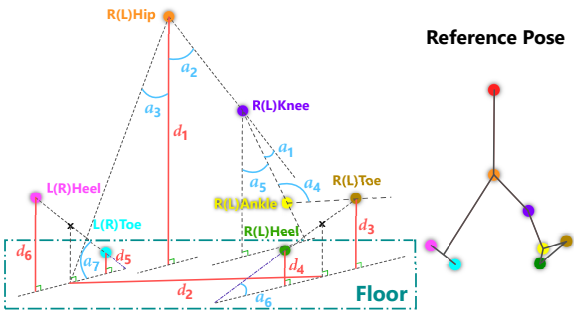

Fig. 2: Visual representation of distances d_i and angles a_i that are used to define gait parameters in Table I. 'x' denotes the 3D position at a segment centroid of two relevant joints represented by the same colour-coded cycles as in Fig. 1. Note that in this figure, the camera is placed at the subject's right side. Thus R (right) ankle is captured while L (left) ankle is only tracked when the camera is placed on the left side.

A. Gait Pattern Extraction

Gait pattern, *i.e.*, the sequence of limb movements during walking, can be characterized by the defined gait parameters (see Table I). While gait patterns are periodic as shown in Fig. 1, they vary among subjects due to differences in age, activity type, gender, proportion and health status. We define a gait pattern as the joint movement, i.e., changes in $\lambda_1, \cdots, \lambda_{12}$, occurring on a complete normalized gait cycle, invariant of walking speed and direction.

Gait cycle detection: Before we characterize the gait patterns, each gait cycle is split based on heel strike event detection. The heel strike event occurs when the foot makes contact with the ground, heel-first; thus, it can be extracted by detecting change points of distances between heel joint and the ground, via the inflection points search method of [4].

Based on trajectories of joints of interest, obtained, for example, by tracking the markers placed on the joints in the recorded video, as in [4], gait parameters (see Table I) are calculated in each frame. For the j-th gait cycle, let $V_j^{\lambda_i}(f)$, $i = 1, \cdots, 12$, denote the value of gait parameter λ_i in Frame f given a gait parameter sequence. For example, $V_1^{\lambda_2}(10)$ denotes the value of knee angle ratio of gait cycle 1 at Frame 10 within a gait parameter sequence.

To *extract distinctive gait patterns* after gait cycle detection, we adopt the following three steps: (Step1) resample complete gait cycles into a fixed length of L samples, called standardized gait patterns; (Step2) cluster similar gait patterns into groups per gait parameter using density-based spatial clustering of applications with noise (E-DBSCAN) [42]; (Step3) generalize gait patterns through DTW-Barycenter Averaging (DBA) [43]. We describe these three steps one by one, next.

Step1: Gait pattern standardization. First, we adopt gait pattern standardization to mitigate the influence of varying gait speed. Fig. 4 shows an example of gait phase periodicity for our proposed 12 gait parameters. Since, in general, $V_j^{\lambda_i}$ varies for different gait cycles, *i.e.*, different j, we adopt *2nd-cubic Bezier curve interpolation* to resample $V_j^{\lambda_i}$ to a fixed length of L samples leading to a standardized gait parameter curve $S_j^{\lambda_i}(x = 1, \cdots L)$, denoted in the following by S_j^z, $z \in \{\lambda_1, \cdots, \lambda_{12}\}$.

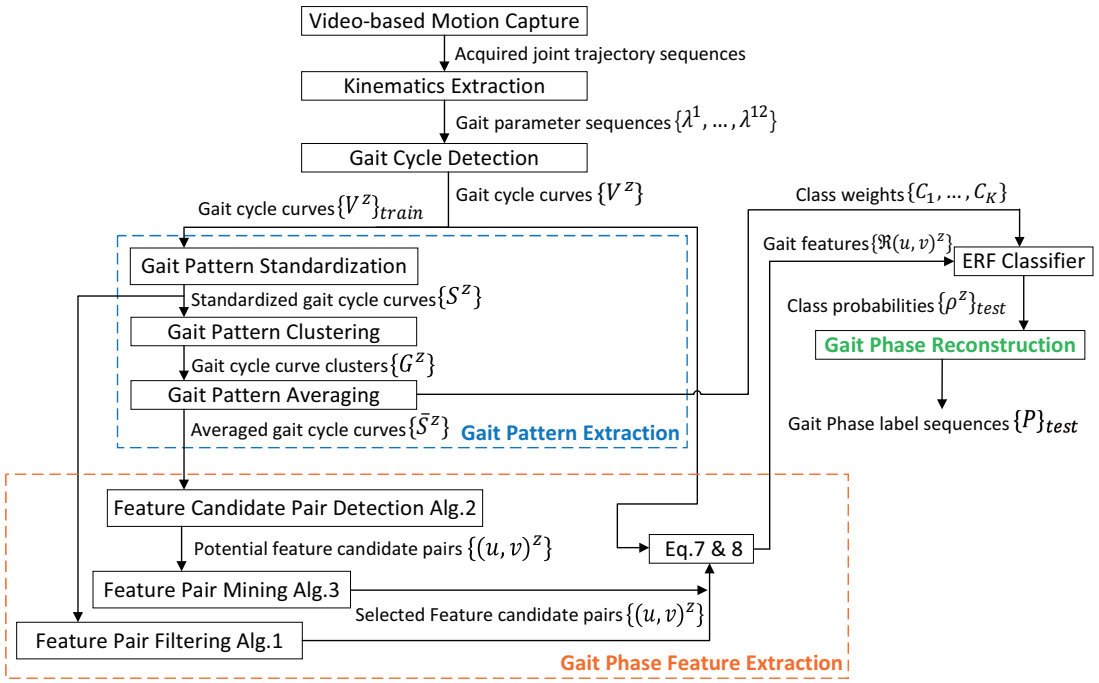

Fig. 3: Overall system diagram. Note that either Alg. 1 (proposed Filtering method) or Algs. 2 and 3 (proposed Optimized method) are used.

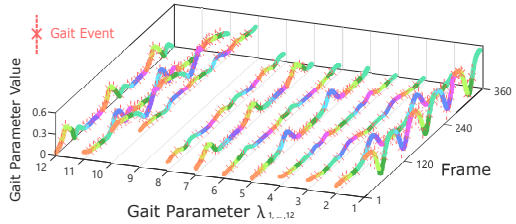

Fig. 4: An example of the gait pattern representation (360 frames are used in this gait sequence). Note that all 12 gait parameter curves and their gait phase progress jointly characterize gait patterns.

Step2: Gait pattern clustering. Clustering analysis is usually used to group similar gait patterns in order to further perform abnormal detection, person recognition, etc. Gait phases are usually ignored in the literature for those tasks. However, similar standardized kinematics might have different gait phases across subjects and this variability is important to capture key gait events. For example, in [44], a wearable sensor-based gait phase detection system uses gait phase duration to classify level-walking and walking upstairs and downstairs by learning a decision tree model. In this paper, we take both gait parameter change and gait phase transition into account to capture the differences between standardized gait patterns. The function $\mathbf{\Phi}^z_{i,j}(x,y)$ that measures the distance between sample x in i-th gait pattern S^z_i and sample y in j-th gait pattern S^z_j is defined as:

$$\mathbf{\Phi}^z_{i,j}(x,y) = \big(S^z_i(x) - S^z_j(y)\big) \cdot \exp\big(\phi(P_{S^z_i(x)}, P_{S^z_j(y)})\big), \tag{1}$$

where

$$\phi(P_a,P_b) = \lceil\frac{K-1}{2}\rceil - \Big|\mathbb{1}\big(\lfloor\frac{2|P_a-P_b|}{K+1}\rfloor \bmod 2 = 0\big)\cdot \lceil\frac{K-1}{2}\rceil - \Big(|P_a-P_b| \bmod \lceil\frac{K+1}{2}\rceil\Big)\Big|. \tag{2}$$

Note that $P_{S^z_i(x)} \in \{1,2,\cdots,K\}$ refers to the corresponding gait phase label at sample x in i-th the gait pattern S^z_i. $\mathbb{1}(p)$ is an operator that returns 1 if a Boolean expression p is true, and 0, otherwise. We use Eq. (1), including both the numerical distance $(S^z_i(x) - S^z_j(y))$ and exponential label distance $\exp\big(\phi(P_{S^z_i(x)}, P_{S^z_j(y)})\big)$, instead of other conventional distance functions that are commonly used in DTW. Furthermore, we employ parameter-free E-DBSCAN to obtain groups of gait patterns via DTW with distance function $\mathbf{\Phi}^z_{i,j}(x,y)$. Each cluster group $G^z = \{\mathring{S}^z_1, \cdots, \mathring{S}^z_M\}$ represents M similar gait patterns.

Step3: Gait pattern averaging. After E-DBSCAN clustering, we employ DBA algorithm to extract labeled gait parameter curves to represent distinctive gait patterns. We define the averaged curve $\bar{S}^z$ as a curve that averages all curves in the same cluster and is initialized as the curve with minimum Euclidean distance compared with all other curves. Then, for each cluster, distance function $\mathbf{\Psi}^z_m(x,y)$ is the conventional DTW distance function between the sample x in the averaged curve $\bar{S}^z$ and the sample y at any $\mathring{S}^z_m$ in the m-th cluster, $m = 1, \cdots, M$, given by:

$$\mathbf{\Psi}^z_m(x,y) = \exp\Big(-\bar{\omega}^z_m(x,y)\cdot\big(\mathring{S}^z_m(y) - \bar{S}^z(x)\big)\Big), \tag{3}$$

where a shared weight $\bar{\omega}^z_m(x,y)$ is computed as:

$$\bar{\omega}^z_m(x,y) = \frac{\sum_{k=1}^{K}\Big(\mathbb{1}\big(P_{\mathring{S}^z_m(y)}=k\big)\Big)}{\sum_{k=1}^{K}\sum_{y'=1}^{m_y}\Big(\mathbb{1}\big(P_{\mathring{S}^z_m(y')}=k\big)\Big)}\cdot \exp\Big(\phi\big(P_{\bar{S}^z(x)}, P_{\mathring{S}^z_m(y)}\big)\Big). \tag{4}$$

Note that we measure the number of occurrences of gait phase label k, in all $\mathring{S}^z_{1,\dots,M}$ within the same cluster. We also use the exponential label distance to weight the distance

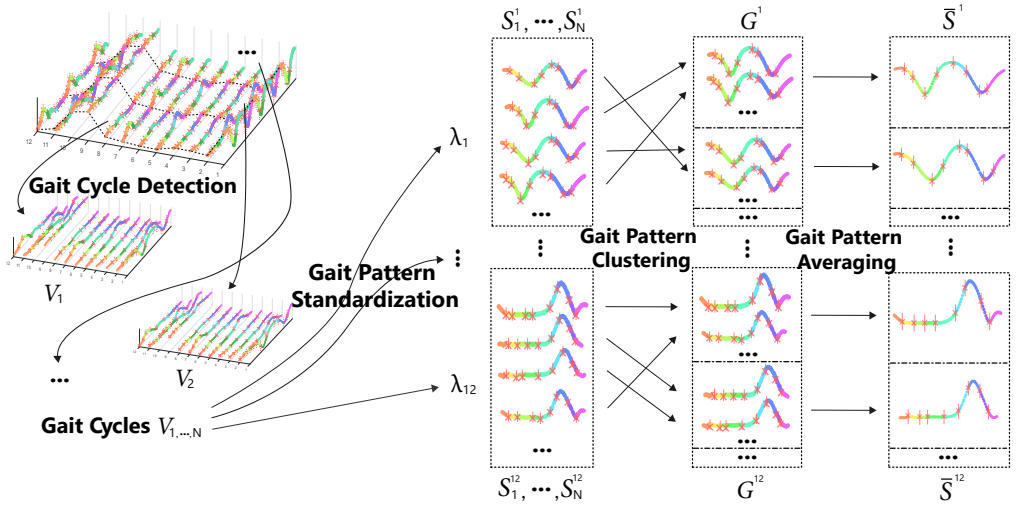

Fig. 5: Proposed gait pattern extraction, showing steps from left to right: extraction of gait parameter curves V^{λ_i} via gait cycle detection, then standardizing to S^z (Step1), clustering to G^z (Step2) and averaging as $\bar{S}^z$ (Step3).

$\mathring{S}^z_m(y) - \bar{S}^z(x)$ in Eq. (3). We adopt these two measurements to tradeoff: 1) between $\mathring{S}^z_m(y)$ and cluster mean $\bar{S}^z(x)$; 2) among gait phase labels $P_{\bar{S}^z(x)}$ and $P_{\mathring{S}^z_m(y)}$ for better gait pattern generalization.

To obtain the corresponding gait phase label for each sample at the averaged curve $\bar{S}^z$, we first extract all warped paths between the averaged curve and all curves within a cluster via DTW using distance function Eq. (3). Then, we adopt min-max standardization to limit the path cost at each iteration of DBA. Finally, for each sample x, we measure the summed path cost for each possible label across all warped paths at all iterations and then label sample x as the label with minimum-sum path cost. Therefore, both numerical gait parameter value averaging and gait phase progress averaging are taken into account to obtain a good representation of a set of similar gait patterns.

Fig. 5 shows the overall gait pattern extraction process, with appropriate notation. The resulting series of the standardized curves $\bar{S}^z$ are obtained by performing Steps 1-3 for the purpose of data reduction from a large number of S^z curves. For example, we obtain 2 to 7 clusters in $\bar{S}^z$ for each $z \in \{\lambda_1, \cdots, \lambda_{12}\}$, from a total 205 gait patterns S^z during our experiments.

Note that the class weights $C_{k=1,\cdots,K}$ of each gait phase label are further used during classification to address the class imbalance problem via:

$$C_{k=1,\cdots,K} = \frac{12 \cdot L}{\sum_{z=\lambda_1}^{\lambda_{12}} \sum_{l=1}^{L} \left(\mathbb{1}\left(P_{\bar{S}^z(l)} = k\right)\right)}. \tag{5}$$

B. Gait Phase Feature Extraction

Recall that our task is to classify each frame into one of $K = 9$ gait phases. Two main challenges in this multi-class gait phase classification problem are: (1) partial gait parameter curves due to both incomplete gait cycles and missing values due to the occlusion at joints of interest; (2) subject-sensitive gait patterns with varying gait speeds. In order to address these two challenges, inspired by image classification study [45], we introduce the feature candidate pair and sliding window, to mitigate feature alignment, and random decision forest to handle the missing observations due to partial gait cycles.

In general, feature engineering of gait phase classification comprises: (1) Feature Alignment: extracting time-varying features to solve temporal misalignment that causes feature mismatch, such that extracted abstract features are invariant to gait speed and subject-dependent patterns. (2) Feature Mining: mine distinct feature pairs, reducing feature candidate pair set without sacrificing the performance.

1) Proposed Feature Alignment: To extract the transition information of adjacent gait phases, we use gait pattern curves $\bar{S}^z$ obtained as explained in Sec.III-A (Step 3). We slide a window across the frames, in order to extract time-varying and linearly separable temporal features. The extracted full gait cycle length L is set to be the length of the sliding window. Since some gait cycles are incomplete, we estimate the length of each incomplete gait cycle from full gait cycles based on the speed of the hip marker.

Considering the continuity of two adjacent gait pattern curves, we make $\bar{S}^z$ periodic such that its starting frame is connected to the end frame (see Fig. 6) of its replica by:

$$\tilde{S}^z = \frac{\bar{S}^z(1) + \bar{S}^z(L)}{2}$$

$$\bar{S}^z(x) = \begin{cases} \bar{S}^z(x+L) & x \in 1-0.5L, \cdots, 1 \\ \bar{S}^z(x) + \frac{(\tau L - x)(\tilde{S}^z - \bar{S}^z(1))}{\tau L} & x \in 1, \cdots, \tau L \\ \bar{S}^z(x) & x \in \tau L, \cdots, L - \tau L \\ \bar{S}^z(x) + \frac{(\tau L + x - L)(\tilde{S}^z - \bar{S}^z(L))}{\tau L} & x \in L - \tau L, \cdots, L \\ \bar{S}^z(x-L) & x \in L, \cdots, 1.5L, \end{cases} \tag{6}$$

where we set the continuous edge ratio to $\tau = 0.1$ to make sure that the boundaries between adjacent $\bar{S}^z$'s are smooth. For example, the target point $\bar{S}^z(x)$ at any position x of the gait pattern curve is smooth within an L-length sliding window. Since the standardized curve is obtained by its replicas, we keep the gait phase labels the same as the ones extracted during

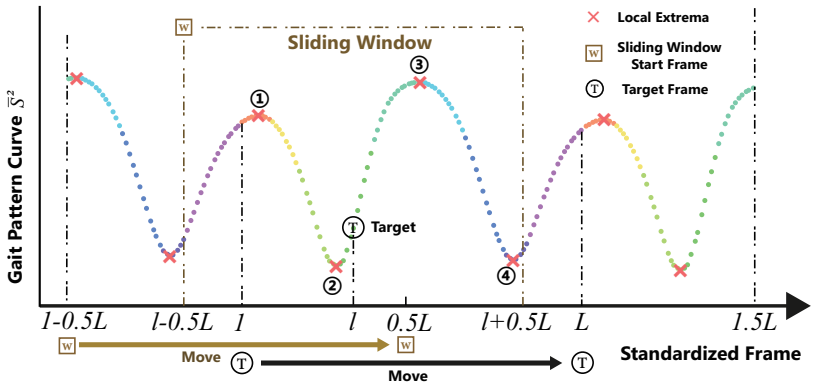

Fig. 6: Feature candidate pair extraction in standardized frame domain: for target frame $l \in 1, \ldots, L$, different candidate pairs (u, v), that link frames with local extrema, give different value responses in terms of knee angle ratio λ_2. Same frames with local extrema give different (u, v) values for different target frames from 1 to L and its corresponding L-length sliding window starting from frame $1 - 0.5L$ to $0.5L$.

gait cycle averaging. The standardized time stamp at sample x becomes x/L given a fixed length of L samples of a sliding window, centered at x. In order to address misalignment of features, we introduce feature pair (u, v), where u and v take values in the range $[-0.5, 0.5]$ with a minimum resolution step of $1/L$ to describe any two points near x within the sliding window.

To capture the time-varying features for each sample x within an L-length sliding window on the periodic gait pattern curve $\bar{S}^z(x)$, $z \in \{\lambda_1, \cdots, \lambda_{12}\}$, given a feature pair $(u, v)^z$ for gait parameter z, we calculate the feature value as:

$$\Re_l(u, v) = \frac{\|\bar{S}^z(l + uL) - \bar{S}^z(l + vL)\|}{|u - v|}, \tag{7}$$

where we set $L = 100$ during the training stage for all extracted gait patterns $\bar{S}^z$. As a result, we have $\binom{L}{2} = 4950$ feature candidate pairs for each gait parameter, including $(-0.5, -0.49), (-0.5, -0.48), \cdots, (-0.5, 0.5), (-0.49, -0.48), \cdots, (0.49, 0.5)$ as $(u, v)^z$.

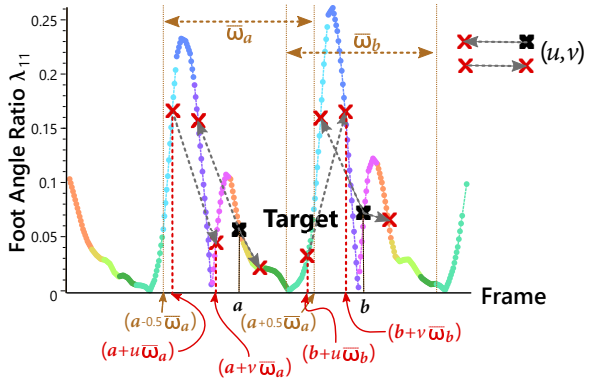

Fig. 7: Feature pairs (u, v) at target frames a and b. We can observe different values of foot angle ratio parameter λ_{11} for the same (u, v) pair at the target frames a and b.

In order to ensure the same time-varying feature statistics for the same (u, v) at the testing stage, we introduce the following equation to calculate the feature value:

$$\Re_f(u, v) = \frac{\|V^z(f + u\varpi_f) - V^z(f + v\varpi_f)\|}{|t^z(f + u\varpi_f) - t^z(f + v\varpi_f)|}, \tag{8}$$

In the above equation, for a frame f, $t(f + u\varpi_f)$ denotes the u shifted timestamp from the time at which target frame f is acquired given a ϖ_f-frame long sliding window centered at frame f. Thus, a gait sequence sampled from frame $f + u\varpi_f$ to frame $f + v\varpi_f$ represents a complete gait cycle. We demonstrate the feature pairs (u, v) in the time domain in Fig. 7.

2) Proposed Feature Mining: To reduce the size of the feature candidate pairs set, which we denote Ω, we select only the most informative pairs. We employ enhanced random decision forests (ERF) [46] to mitigate the issue of missing feature values that often occur near the boundary of the gait parameter curves.

Two mining approaches are proposed to extract the feature pairs: (1) filtering method: enumerate all (u, v) link combinations in terms of standardized length $L = 100$ such that the quality of $12\binom{L}{2}$ feature candidate pairs for all gait parameters is evaluated during mining; (2) optimized method: find all (u, v) link combinations of local extrema frames via Persistence1D [47]. We explain these two approaches and evaluate their performance next.

2.1) **Filtering method (Alg. 1)**: We enumerate all possible (u, v) link combinations of any two frames within an L-length sliding window centered on the target frame. Then we compute the feature values using (7) from the standardized curve $\bar{S}^z$ extracted in Sec. III-A. To ensure the feature values are in the same feature space as the ERF model, we use Gini impurity [48] as evaluation metric to quantify the information content of a feature candidate pair. To avoid overfitting the ERF model, we select $\sum_{z=\lambda_1}^{\lambda_{12}} num^z$ best feature candidates from a total of $12\binom{L}{2}$ candidate pairs, where num^z is the resulting feature count for each gait parameter z. The distinct feature pairs are chosen via the following *Criteria* and Alg. 1: (1) Given a finite set of feature candidate pairs $\{(u, v)^z_\varepsilon\}$ for each gait parameter z, we compute Gini impurity $\mathcal{G}^z_\varepsilon(P_a, P_b)$ (see Alg. 1) to separate samples with gait phase labels P_a and P_b, $a, b \in \{1, \cdots, K\}$ as a measure of quality $Q^z_\varepsilon(P_a, P_b)$ for the ε-th feature candidate pair $(u, v)_\varepsilon$ for gait parameter z. (2) For each z, we calculate the quality Q^z for all possible gait phase label combinations $\mathcal{H} = \{(P_1, P_2), (P_1, P_3), \cdots, (P_2, P_3), \cdots, (P_{K-1}, P_K)\}$

by:

$$Q^z = \sum_{\varepsilon=1}^{\binom{L}{2}} \sum_{H \in \mathcal{H}} Q_\varepsilon^z(H). \tag{9}$$

(3) To maximize the total quality, only num^z distinct pairs with the highest quality are considered, where

$$num^z = \frac{\Omega \sum_{\varepsilon=1}^{\binom{L}{2}} \sum_{H \in \mathcal{H}} Q_\varepsilon^z(H)}{\sum_{z=\lambda_1}^{\lambda_{12}} \sum_{H \in \mathcal{H}} Q^z(H)}. \tag{10}$$

Algorithm 1: Feature candidate pair filtering.

Input: Standardized gait pattern curves $\{S^z\}$ (Sec.III-A); Potential feature candidate pair list $\{(u,v)^z\}$;
Output: Feature pair list $\{(u,v)^z\}$;

update standardized gait pattern curves $\{\bar{S}^z\}$ via Eq. 6;
foreach $z \in \{\lambda_1, \cdots, \lambda_{12}\}$ **do**
  foreach $H = (P_a, P_b) \in \mathcal{H}$ **do**
    foreach $(u,v) \in \{(u,v)^z\}$ **do**
      compute $\{\Re_l(u,v)^z\}$ for all $\{\bar{S}^z\}$ via Eq. 7;
      find best split for P_a and P_b that minimizes $\mathcal{G}_\varepsilon^z(P_a,P_b) = 1 - p_{P_a}^2 - p_{P_b}^2$ from $\{\Re_l(u,v)^z\}$;
      update $\{\Re_l(u,v)^z\}$ for all $\{S^z\}$ via Eq. 7;
      update $\mathcal{G}_\varepsilon^z(P_a,P_b)$ using the found best split;
      set $Q_\varepsilon^z(H) = 1 - \mathcal{G}_\varepsilon^z(H)$;
  update Q^z via Eq. 9;
estimate num^z using Eq. 10;
foreach $z \in \{\lambda_1, \cdots, \lambda_{12}\}$ **do**
  sort $\{(u,v)^z\}$ by Q^z;
  keep top num^z in $\{(u,v)^z\}$;

2.2) **Optimized method (Algs. 2 and 3)**: We observe heuristically that most of the distinct feature pairs link two frames where at least one frame has a local extrema value. Given the large time-varying feature value response and availability of a more distinguishable difference for classifying two gait phases, a computationally more efficient way to extract feature candidate pairs (u,v) is to find those pairs that link local extrema.

For a standardized gait pattern curve $\bar{S}^z$, we propose Alg. 2 to extract feature candidate pairs $\{(u,v)^z\}$ where neighbour radius r_z is designed to discard duplicate feature candidate pairs with similar feature values during adjacent frames. A good value for r_z is determined as follows: find the minimum frame length for which any two distinguishable frames are separated, and calculate the median duration $\ddot{T}_{p=1,\cdots,K}$ for each gait phase in the training sets. For example, for our datasets, we observed that $\min(\ddot{T}_{p=1,\cdots,K}) \approx 6\%$ of a gait cycle; therefore, we need to access at least $r_z = 6\%L = 6$ frames to observe two distinguishable frames within the same gait phase. Note that a standardized gait pattern curve $\bar{S}^{z=2}$ for knee angle ratio λ_2 is first extracted as explained in Sec. III-A (see Fig. 6).

As shown in Fig. 6, the standardized frames from 1 to L are used to generate periodic gait pattern curve frames from $1 - 0.5L$ to $1.5L$ using Eq. (6). Next, an L-length sliding window moves from frame $1 - 0.5L$ to $0.5L$ with its corresponding target frame l from 1 to L for full enumeration of standardized gait pattern frames in terms of all potential feature candidate

Algorithm 2: Feature candidate pair detection.

Input: Standardized gait pattern curves $\{\bar{S}^z\}$ (Sec.III-A); Neighbor radius r_z;
Output: Potential feature candidate pair list $\{(u,v)^z\}$;

update standardized gait pattern curves $\{\bar{S}^z\}$ via Eq. 6;
foreach *frame* $l \in \{1, \cdots, L\}$ **do**
  find frames $\{f_{i=1:n}\}$ from frames $l - 0.5L + 1, \cdots, l + 0.5L - 1$ with local extrema and persistences $\beta = \{\beta_{i=1:n}\}$ via [47];
  if $n < 2\sqrt{L}$ **then**
    $\beta_{threshold} = median(\beta)$;
  else
    $\beta_{threshold} = 0$;
  keep frames $\{f_{j=1:\mathring{n}}\}$ with corresponding persistence $\beta_j < \beta_{threshold}$;
  for $U = \{f_{j=1:\mathring{n}}\}$ **do**
    for $V = \{U, \cdots, f_{j=\mathring{n}}\}$ **do**
      if $V - U > r_z$ **then**
        add $(u,v) = (\frac{U-l}{L}, \frac{V-l}{L})$ to feature candidate pair list $\{(u,v)^z\}$;

Algorithm 3: Feature pair mining.

Input: Potential feature candidate pair list $\{(u,v)_1^z, \cdots, \{(u,v)_{N_z}^z\}$ for each gait parameter $z \in \{\lambda_1, \cdots, \lambda_{12}\}$;
Desired Feature Count Ω;
Output: Feature pair list $\{(u,v)^z\}$ for each gait parameter z;

foreach z *in* $\{\lambda_1, \cdots, \lambda_{12}\}$ **do**
  for $n_z = \{1, \cdots, N_z\}$ **do**
    estimate probability of each candidate pair by $O_{n_z} = \frac{num((u,v)_{n_z})}{num((u,v)_1, \cdots, (u,v)_{N_z})}$;
  let $O_{median^z} = median(\{O_1, \cdots, O_{N_z}\})$;
foreach $(u,v)^z$ *in* $\{(u,v)^z\}$ **do**
  extract gait phase probabilities $\{p_{(u,v)^z}^1, \cdots, p_{(u,v)^z}^K\}$ when $(u,v)^z$ is detected at its target frame l;
foreach z *in* $\{\lambda_1, \cdots, \lambda_{12}\}$ **do**
  set feature count $\Omega^z = \frac{\Omega \cdot O_{median^z}}{\sum_{z=\lambda_1}^{\lambda_{12}} O_{median^z}}$;
  sort feature candidate pairs $\{(u,v)^z\}$ descending by probability O_{n_z};
  group $\{(u,v)^z\}$ by its top 2 gait phase labels with relative high gait phase probabilities;
  add Ω^z $(u,v)^z$ in total of groups to feature pair list $\{(u,v)^z\}$ balancedly based on class weights $C_{k=1,\cdots,K}$;

pairs. For each enumerated frame inside the sliding window, we adopt Persistence1D [47] to extract local extrema and its persistence. In order to filter out those extrema with a relatively small value change, we heuristically use median persistence as threshold. For the remaining frames with local extrema, we link any two frames and compute its normalized offset to the target frame l as feature candidate pair (u, v).

In order to reduce the feature dimension, we propose Alg. 3 to select the potential feature candidate pairs based on their probabilities of occurrence in terms of gait parameter $z \in \{\lambda_1, \cdots, \lambda_{12}\}$. This reduces the number of input feature candidate pairs when globally optimal feature selection criteria is applied, such that significant computation time is saved to maximize the factor of variation between gait phases.

Given a total of $12\binom{L}{2} = 59,400$ feature candidate pairs for all gait parameters $z \in \{\lambda_1, \cdots, \lambda_{12}\}$, the filtering method examines each (u, v) pair by Gini impurity and selects those candidate pairs as per Eq. (10) on all extracted standardized gait patterns $\bar{S}$ within the training samples. Instead of this exhaustive search (Alg. 3), the proposed optimized method first detects potential feature candidate pairs via Alg. 2 where those pairs with low feature response are dropped. Afterwards, the pairs will be selected according to their probabilities of occurrence in the training samples, via Alg. 3. Specifically, our proposed optimized feature extraction methods (Alg. 2 and Alg. 3) decrease the computational complexity of evaluating all feature candidate pairs from $O(12 \cdot \binom{K}{2}\binom{L}{2} \cdot M \log(M))$ (recall that M is the total number of the input standardized curves $S^z \ \forall z \in \lambda_1, \cdots, \lambda_{12})$ to $O(12ML \log(L)) + O(M \sum_{z=\lambda_1}^{\lambda_{12}} N_z)$ compared with the filtering method in Alg. 1.

C. Gait Phase Reconstruction

After the ERF-based frame-wise classification, we re-segment each gait phase period using the proposed gait phase reconstruction approach described next.

Since the gait phase is defined as the period between two adjacent gait events, the gait phase label sequence is a periodic piece-wise smooth signal. The gait phase label sequence obtained by the gait phase classifier, described in the previous subsection, sometimes contains classification errors and could even be non-periodic. To improve the final classification result, we use the fact that the labels sequence changes smoothly over time.

In particular, we first locate a gait event as a time-stamp within two adjacent gait phases. Then, we obtain a class probability vector $\rho_f = \{\rho_f^1, \cdots, \rho_f^K\}$, given feature values $\Re^z(u, v)$ for each frame f from the trained ERF model as discussed in Sec. III-B. Since, as shown in Fig. 8, misclassification often occurs near the boundary of two adjacent gait phases, we introduce a correlation coefficient η_f in Eq. (11) to capture similarity between adjacent gait events P_a and P_b, within the frames $f_s, \cdots, f_e$:

$$\eta_f(a, b) = \sum_{i=f_s}^{f} \rho_i^a \sum_{i=f}^{f_e} \rho_i^b - \sum_{i=f_s}^{f} \rho_i^b \sum_{i=f}^{f_e} \rho_i^a. \qquad (11)$$

As shown in Fig. 8, we refine the gait phase labels by detecting the gait event at the moment when the correlation coefficient reaches the global maximum. We keep updating, iteratively, the gait phase labels until convergence.

IV. System Validation

In this section we report our experimental results[2]. Each step proposed in the methodology section is assessed in terms of classification accuracy and complexity in order to show its importance towards the final result. Each step is benchmarked against relevant algorithms in the literature.

Joint trajectories are obtained using the MS Kinect v2-based motion capture system of [4], which does not rely on Kinect skeleton data and shows close accuracy to commercial 12-camera VICON system [14]. Note that, our proposed algorithms are applicable to other motion capture systems, requiring only 10 standard joint trajectories defined in the widely used Plug-in-Gait model [49] for gait assessment as input. However, the overall gait phase classification performance will depend on the tracking accuracy of the employed motion capture system.

9 stroke survivors and 6 healthy volunteers were asked to walk for 6 meters while the middle 4-meter motion is recorded at two different rehabilitation sessions. All 15 participants read the participant information sheets and completed the consent forms before data collection The study covered in this paper is conducted with ethical approval from both National Health Service (NHS) and University of Strathclyde Ethics Committee (UEC). As per similar experiments in [12], [27], [30] where 15 subjects, 10 subjects (116 strides), and 25 strides are used, respectively, in our experiments, 126 records (613 strides) are captured from 15 subjects with various walking speeds, directions and patterns. We acquire trajectories of hip, knee, ankle, toes, heels joints during the 4-meter walking test. To form ground-truth for classifying $K = 9$ gait events, we first manually identify all frames that separate adjacent gait phases by finding specified gait poses defined in [41]. Then, the timestamps are carefully refined by pose template matching on interpolated gait parameters (see Table I). Finally, we assign a gait phase label to each frame by slicing the extracted gait events.

To evaluate the classification performance, we use classification accuracy (ACC), average receiver operating characteristic (ROC) curve [22], [50] and area under the curve (AUC) [50] as evaluation metrics. ACC is defined as the percentage of correctly classified frames across all sequences. The average ROC is plotted by measuring the true positive rate (TPR) and false positive rate (FPR) based on one-versus-all binary classification across all sequences and all gait phases. For each gait phase k, the corresponding TPR measures the proportion of frames at phase k that are correctly classified. FPR calculates the proportion of frames not at gait phase k that are incorrectly classified as gait phase k frames. We divide our dataset into training and testing set, where the training set comprises stroke survivors 1-5 and healthy volunteers 1-3, and testing

[2]Code to reproduce our experiments is available at https://github.com/yemx21/GaitPhaseClassification

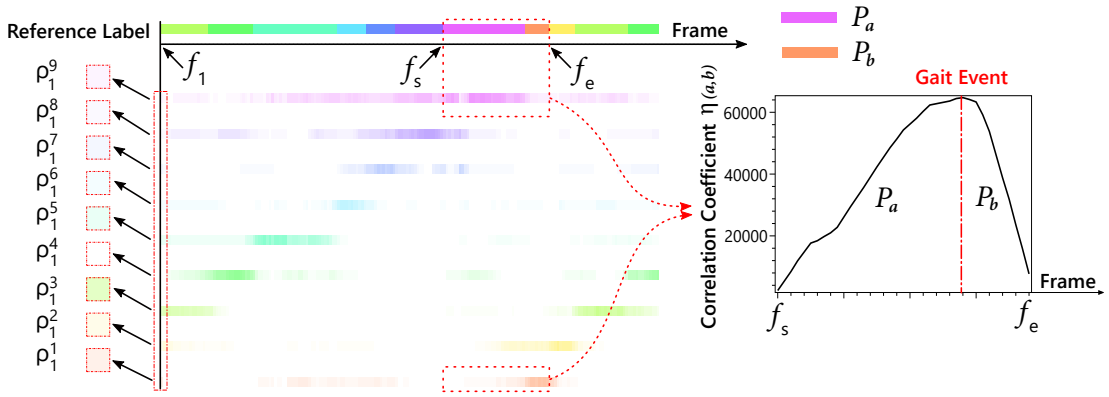

Fig. 8: Class probability representation in a complete gait cycle and correlation coefficient for adjacent gait phase periods. Unique color is assigned to each gait phase label where low transparency means low probability of the label occurrence at specified frame.

set comprises stroke survivors 6-9 and healthy volunteers 4-6. Grid search is adopted to tune hyper-parameters.

We evaluate the importance of different steps in the proposed system (Fig. 3) in the following way. First, to justify the selection of ERF as classifier, we tested One vs One multi-class SVM and classical two-layer softmax NN against ERF. To test the reliability and validity of the defined 12 gait parameters at frame-wise level, we feed to the classifiers normalized trajectories and compare performance when the proposed 12 gait parameters are used as features instead. To assess the value of gait phase reconstruction, we test schemes with and without this step.

We group our evaluations and benchmarking as follows: (1) Schemes denoted with SVM[1], NN[1], and ERF[1] use normalized joint trajectories as input (as discussed in [51]), *e.g.*, acquired joint trajectory sequences from the output of video-based motion capture are fed directly into the classifier (after normalization); (2) Schemes denoted with the classifier name without any superscripts, *e.g.*, SVM, NN, and ERF, perform Kinematics Extraction and the proposed Gait Cycle Detection (see Fig. 3) and feed the resulting gait cycle curves $V^z(\lambda)$ directly into the classifier; (3) The scheme denoted by NARX-NN uses the state-of-the-art NARX-NN model [30] (we use Matlab 2016a Neural Time Series toolbox) on the standardized gait cycle curves S^z; since the NARX-NN performs worse when data with incomplete gait cycle is used, we only use frames with complete gait cycles. (4) NARX-NN[2], **Filtering**[2] (Alg. 1) and **Optimized**[2] (Alg. 2 and Alg. 3), denote NARX-NN, the proposed system with feature candidate pair filtering, and the proposed system with feature candidate pair detection and feature pair mining, respectively, without the Gait Phase Reconstruction block. (5) The proposed **Filtering** and **Optimized** schemes with all the steps, *e.g.*, including Gait Phase Reconstruction.

Note that for the NARX-NN model, given a regression result Γ_f at frame f [30], its corresponding gait phase label is $P(f) = \mathbb{1}\{\lfloor\Gamma_f\rfloor \mod K = 0\} \cdot K + \lfloor\Gamma_f\rfloor \mod K$, where $\mathbb{1}$ all ones matrix, $\lfloor\cdot\rfloor$ returns the first smaller integer, and $\mod$ returns a remainder, and the corresponding class probability is calculated by:

$$\rho_f^k = \begin{cases} 1 - |k - \Gamma_f| & k \in \{\lfloor\Gamma_f\rfloor, \lceil\Gamma_f\rceil\} \\ 0 & k \notin \{\lfloor\Gamma_f\rfloor, \lceil\Gamma_f\rceil\}. \end{cases} \quad (12)$$

Levenberg-Marquardt method is used to train a two-layered NARX-NN. In the experiments, data is randomly spit into 70% data used for training, 15 used for validation and 15% used for testing. All 9×10 trained networks are evaluated for the 15% testing data for time delays from 1-9 frames with 10 network per time delay. The mean ACC of the NARX-NN model is evaluated in [1], showing that the best result is obtained for an input time delay of 8 frames; this will be used in the following experiments.

For our proposed gait phase classification system, we randomly choose approximately 80% of the training data as training set and the rest as validation to tune hyper parameters. Note that both sets include frames with incomplete gait cycles, to match the training configurations using NARX-NN. We train 20 sets of ERF models with depth 20 and 30 trees to evaluate the ACC for various sizes of feature candidate pair set, Ω.

The results are shown in Table II and Fig. 9. The advantage of an improved characterized gait presentation using the proposed kinematics extraction method can be observed from Table II and Fig. 9. The methods that apply the proposed Kinematics Extraction (SVM, NN, ERF) show a significant ACC and AUC improvement over feeding normalized joint trajectories to the classifiers (schemes with superscript [1]). Indeed, SVM consistently and significantly outperforms SVM[1] with both higher ACC and AUC, while ERF, NN, ERF[1] and NN[1] show similar results, but all being significantly worse than the proposed (Filtering and Optimized) methods. NARX-NN generally outperforms SVM and NN classifiers, but it is consistently outperformed by the proposed method.

NARX-NN[2], Filtering[2] and Optimized[2] algorithms all significantly outperform the ERF classifier. When comparing NARX-NN[2], Filtering[2] and Optimized[2] algorithms with NARX-NN, Filtering and Optimized algorithms, we can observe that the proposed Gait Phase Reconstruction method used in the latter algorithms leads to significant ACC improvement, of up to 15%. Finally, we observe that the proposed

TABLE II: Best performance (in terms of ACC in [%]) for all $K = 9$ gait phase classes for all tested schemes.

Gait Phase Class	1	2	3	4	5	6	7	8	9	Mean
SVM[1]	52.3	38.7	72.8	55.2	66.3	68.7	86.4	83.6	75.7	67.7
NN[1]	62.0	49.1	80.8	80.6	78.7	81.6	90.8	87.4	80.8	77.7
ERF[1]	51.6	54.7	83.3	78.6	73.1	70.6	92.2	83.2	71.2	74.1
SVM	81.4	61.9	74.1	66.4	79.6	87.5	88.5	87.9	82.1	79.3
NN	79.7	60.5	77.3	75.6	80.4	83.7	87.6	89.9	79.6	80.0
ERF	69.7	61.9	78.8	75.0	80.9	84.2	88.5	89.2	77.3	81.1
NARX-NN[2]	79.7	82.4	81.8	80.4	79.4	82.9	84.5	87.3	86.2	83.3
Filtering[2]	81.7	85.5	82.4	83.2	82.1	84.3	82.0	87.8	86.3	84.5
Optimized[2]	83.1	87.2	86.7	85.3	84.3	85.2	83.2	88.5	86.3	86.1
NARX-NN	82.3	91.2	90.3	90.3	81.9	88.2	92.9	88.5	89.1	88.8
Filtering	98.2	99.1	99.1	98.6	97.6	**98.7**	98.4	**98.2**	97.8	98.9
Optimized	**98.5**	**99.3**	**99.2**	**98.9**	**98.0**	**98.7**	**98.7**	**98.2**	**98.0**	**99.2**

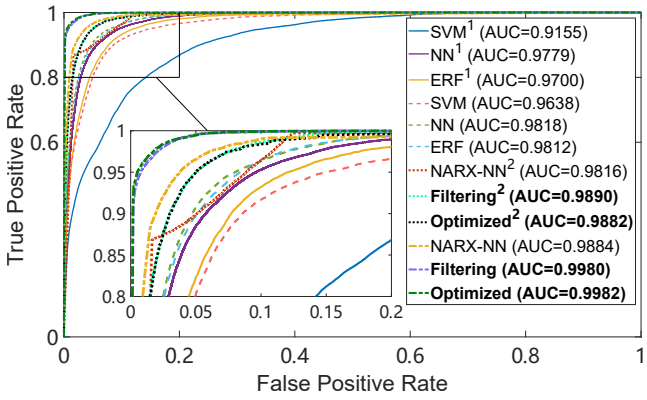

Fig. 9: Average ROC curves for all tested schemes.

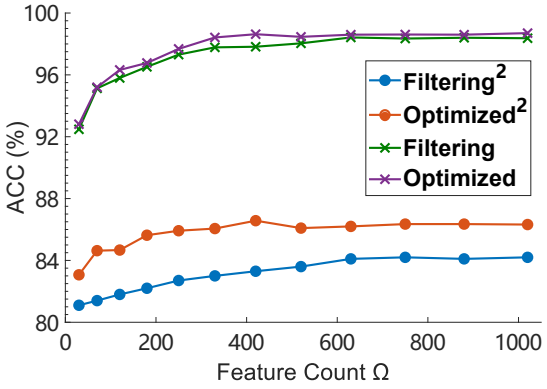

Fig. 11: Testing performance (ACC) of the four mining methods before and after gait phase reconstruction using ERF.

Optimized and Filtering methods consistently outperform all prior benchmarks.

Next, in Figs. 10 and 11, we compare more closely the two proposed approaches of feature candidate pair mining, Filtering and Optimized. Since the Filtering method enumerates $\binom{L}{2}$ feature candidate pairs for each gait parameter during feature extraction and mining, we can observe that the ERF model is over-fitted when the feature number reaches approximately 630 with the highest final ACC. On the other hand, for the Optimized method that extracts all feature candidate pairs with local extrema linkage, a smaller feature number can achieve higher ACC. Both validating and testing ACCs at the classifier stage are improved by using the optimized feature candidate pair detection and mining method (Algs.2 and 3).

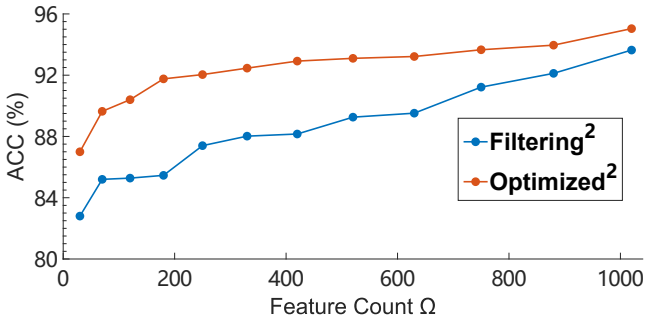

Fig. 10: Validating performance (ACC) of the two mining methods during classifier stage using ERF.

To evaluate the **computational efficiency**, we define $Speedup = T_{baseline}/T_{evaluate}$, where $T_{baseline}$ is the average total execution time required to train a classifier using the proposed filtering method, while $T_{evaluate}$ refers to the average total execution time for obtaining the results using the evaluated methods (*e.g.*, Alg. 2 + Alg. 3, or NARX-NN). For instance, the $Speedup$ of optimized mining approach during testing equals to the execution time ratio between optimized and filtering mining approach on testing sets using corresponding trained models. The experiments were performed on an Intel i7-4710HQ 2.5GHz CPU, Windows10 OS, implemented using Visual C++ and Matlab 2016a.

Fig. 12 confirms that there is a large computation performance boost using the proposed optimized feature candidate mining approach compared to the filtering method. The reduction in computation time comes from detecting feature candidate pairs with local extrema linkage before selection instead of using Gini impurity calculations.

Since the filtering feature extraction firstly computes qualities of all potential feature candidate pairs for each gait parameter, the computation time of feature extraction does not depend on the feature count Ω. Unlike the filtering approach, the proposed optimized feature extraction firstly detects the feature candidate pairs by detecting local extrema, and then narrows the potential candidates based on probabilities of different gait phase label combinations. This filters out potential feature candidate pairs before the feature selection process. We

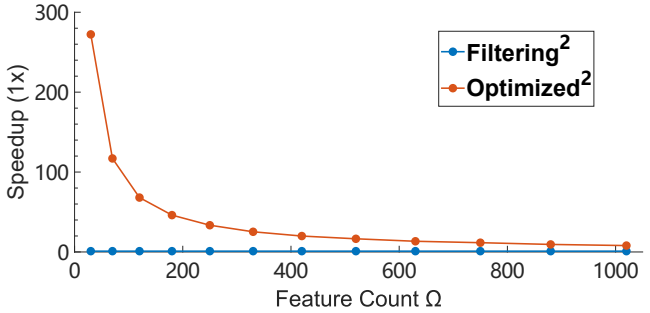

Fig. 12: Computation performance (1x times of baseline) of the proposed feature extraction methods: the base line computation time is measured using the filtering method with 1020 feature candidate pairs.

also observe a slight improvement in the final ACC and smaller best ACC feature count after the gait phase reconstruction process, as shown in Fig. 11.

Finally, we compare computational efficiency with the NARX-NN model-based method, where for fair comparison, we use the frames with full gait cycles only. The parameters for each method are set to achieve their best final ACCs. We report *Speedup* of the three approaches in Table III. Note that the corresponding best ACC results for these three methods are shown in Table II. It can be seen that Optimized method is the fastest method, 2 to 6 times faster than the NARX-NN model.

TABLE III: *Speedup* (1x times of baseline) of the proposed feature extraction methods and NARX-NN model corresponding to their highest ACC.

Method	Filtering	Optimized	NARX-NN
Train	1	18.2	3.1
Test	253	1286	1110

V. Conclusion

We propose a gait phase classification system suitable for clinical diagnostics during rehabilitation programmes. Specifically, we propose gait pattern extraction based on 12 defined gait parameters to help feature extraction from a noisy large-scale dataset via clustering and averaging using proposed distance functions. We mitigate feature misalignment by introducing gait pattern standardization. Optimized feature candidate pair detection and mining algorithms (Algs. 2, 3) are proposed to reduce the computational complexity without sacrificing the classification performance. In order to refine the gait phase labels, we developed a classification approach to relocate the gait events between adjacent gait phases. The overall proposed, frame-wise, multi-channel, time-series classification algorithm demonstrates a significant accuracy and computational complexity improvement with respect to the state-of-the-art NARX-NN model.

VI. Acknowledgment

This project has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No 734331. The University of Strathclyde gratefully acknowledges the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

References

[1] M. Ye *et al.*, "Gait phase classification for in-home gait assessment," in *2017 IEEE International Conference on Multimedia and Expo (ICME)*, July 2017, pp. 1524–1529.
[2] C. Yang *et al.*, "Autonomous gait event detection with portable single-camera gait kinematics analysis system," *Journal of Sensors*, Dec 2015.
[3] C. Yang, U. Ugbolue, D. McNicol, V. Stankovic, L. Stankovic, A. Kerr, B. Carse, K. Kaliarntas, and P. Rowe, "Automation enhancement and accuracy investigation of a portable single-camera gait analysis system," *IET Science, Measurement and Technology*, pp. 1–9, 2019.
[4] M. Ye *et al.*, "A depth camera motion analysis framework for tele-rehabilitation: Motion capture and person-centric kinematics analysis," *IEEE JSTSP*, vol. 10, pp. 877 – 887, Aug. 2016.
[5] B. J. E. Misgeld et al., "Body-sensor-network-based spasticity detection," *IEEE JBHI*, vol. 20, no. 3, pp. 748–755, May 2016.
[6] C. Zhang *et al.*, "Classification of eeg signals using multiple gait features based on small-world neural network," in *Int. Conf. Ubiquitous Robots and Ambient Intelligence*, Aug 2016, pp. 61–66.
[7] Q. T. Ly *et al.*, "Detection of gait initiation failure in parkinson's disease patients using eeg signals," in *2016 38th Annual Int. Conf. the IEEE EMBC*, Aug 2016, pp. 1599–1602.
[8] Y. D. Xuan et al., "Elderly gait analysis and assessment based on body area network," in *GreenCom IEEE Int. Conf. Cyber, Physical and Social Computing and Communications*, Aug 2013, pp. 1729–1733.
[9] M. Derawi *et al.*, "Fusion of gait and ecg for biometric user authentication," in *Int. Conf. BIOSIG*, Sep 2014, pp. 1–4.
[10] "M3d," http://www.tecgihan.co.jp/en/products/m3d_forceplate, Jan 2019.
[11] J.-T. Zhang *et al.*, "Concurrent validation of xsens mvn measurement of lower limb joint angular kinematics," *Physiological Measurement*, vol. 34, no. 8, p. N63, 2013.
[12] J.-Y. Jung *et al.*, "A neural network-based gait phase classification method using sensors equipped on lower limb exoskeleton robots," *Sensors*, vol. 15, no. 11, pp. 27 738–27 759, Oct 2015.
[13] J. Taborri *et al.*, "A novel hmm distributed classifier for the detection of gait phases by means of a wearable inertial sensor network," *Sensors*, vol. 14, no. 9, pp. 16 212–16 234, Sep 2014.
[14] VICON, "Motion capture," http://www.vicon.com, Jan 2019.
[15] Qualisys, "Motion analysis," http://www.qualisys.com, Jan 2019.
[16] C. Yang *et al.*, "Human upper limb motion analysis for post-stroke impairment assessment using video analytics," *IEEE Access*, vol. 4, pp. 650 – 659, Feb 2016.
[17] A. Leu *et al.*, "A robust markerless vision-based human gait analysis system," in *IEEE SACI*, May 2011, pp. 415–420.
[18] Y.-R. Li *et al.*, "A gait analysis system using two cameras with orthogonal view," in *Multimedia Technology (ICMT), 2011 Int. Conf.*, July 2011, pp. 2841–2844.
[19] "Developing with kinect for windows," https://developer.microsoft.com/en-us/windows/kinect, Jan 2019.
[20] E. Gianaria *et al.*, "Kinect-based gait analysis for automatic frailty syndrome assessment," in *IEEE ICIP*, Sep 2016, pp. 1314–1318.
[21] D. Podsiadlo *et al.*, "The timed " up & go": A test of basic functional mobility for frail elderly persons," *Journal of the American Geriatrics Society*, vol. 39, no. 2, pp. 142–148, 1991.
[22] Q. Zou, L. Ni, Q. Wang, Q. Li, and S. Wang, "Robust gait recognition by integrating inertial and rgbd sensors," *IEEE transactions on cybernetics*, vol. 48, no. 4, pp. 1136–1150, 2018.
[23] C. Yang, G. Cheung, V. Stankovic, K. Chan, and N. Ono, "Sleep apnea detection via depth video and audio feature learning," *IEEE Transactions on Multimedia*, vol. 19, no. 4, pp. 822–835, April 2017.
[24] C. Yang, G. Cheung, and V. Stankovic, "Estimating heart rate and rhythm via 3d motion tracking in depth video," *IEEE Transactions on Multimedia*, vol. 19, no. 7, pp. 1625–1636, July 2017.

[25] G. Paolini *et al.*, "Validation of a method for real time foot position and orientation tracking with microsoft kinect technology for use in virtual reality and treadmill based gait training programs," *IEEE TNSRE*, vol. 22, no. 5, pp. 997–1002, Sep 2014.

[26] R. A. Clark *et al.*, "Validity of the microsoft kinect for assessment of postural control," *Gait & Posture*, vol. 36, no. 3, pp. 372–377, 3 2012.

[27] Y. Yang *et al.*, "Reliability and validity of kinect rgb-d sensor for assessing standing balance," *IEEE Sensors Journal*, vol. 14, no. 5, pp. 1633–1638, May 2014.

[28] Z. Liu, Z. Lin, X. Wei, and S. Chan, "A new model-based method for multi-view human body tracking and its application to view transfer in image-based rendering," *IEEE Transactions on Multimedia*, vol. 20, no. 6, pp. 1321–1334, June 2018.

[29] D. Peters *et al.*, "Assessing the reliability and validity of a shorter walk test compared with the 10-meter walk test for measurements of gait speed in healthy, older adults," *Journal of geriatric physical therapy*, vol. 36, pp. 24–30, March 2013.

[30] C. Galván-Duque *et al.*, "Comparison between classical and intelligent identification systems for classification of gait events," *J. Control Science and Engineering*, vol. 1, pp. 21–34, 2015.

[31] R. Vemulapalli *et al.*, "Human action recognition by representing 3d skeletons as points in a lie group," in *IEEE CVPR*, June 2014, pp. 588–595.

[32] R. Vemulapalli *et al.*, "Rolling rotations for recognizing human actions from 3d skeletal data," in *IEEE CVPR*, June 2016, pp. 4471–4479.

[33] S. Zhang *et al.*, "A review on human activity recognition using vision-based method," *Journal of Healthcare Engineering*, July 2017.

[34] Z. Liu and S. Sarkar, "Improved gait recognition by gait dynamics normalization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 6, pp. 863–876, 2006.

[35] J. Zhao *et al.*, "Gait assessment using the kinect rgb-d sensor," in *IEEE EMBC*, Aug 2015, pp. 6679–6683.

[36] F. Gholami *et al.*, "A microsoft kinect-based point-of-care gait assessment framework for multiple sclerosis patients," *IEEE JBHI*, vol. PP, no. 99, pp. 1–1, 2016.

[37] Q. Wang *et al.*, "Unsupervised temporal segmentation of repetitive human actions based on kinematic modeling and frequency analysis," in *3DV 2015*, Nov 2015, pp. 562–570.

[38] J. D. Farah, N. Baddour, and E. D. Lemaire, "Gait phase detection from thigh kinematics using machine learning techniques," in *2017 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, May 2017, pp. 263–268.

[39] C. wei Hsu, C. chung Chang, and C. jen Lin, "A practical guide to support vector classification," http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf, 2010.

[40] S. Mulroy *et al.*, "Use of cluster analysis for gait pattern classification of patients in the early and late recovery phases following stroke," *Gait & Posture*, vol. 18, no. 1, pp. 114–125, 2003.

[41] W. R. Frontera, Ed., *DeLisa's Physical Medicine & Rehabilitation: Principles and Practice*, 5th ed. Lippincott Williams & Wilkins, 2010, vol. 1, ch. 5, pp. 121–137.

[42] P. Sharma *et al.*, *Efficient Density-Based Clustering Using Automatic Parameter Detection*. Springer, 2016, vol. 1, pp. 433–441.

[43] F. Petitjean *et al.*, "A global averaging method for dynamic time warping, with applications to clustering," *Pattern Recognition*, vol. 44, no. 3, pp. 678 – 693, 2011.

[44] J. S. Wang *et al.*, "Walking pattern classification and walking distance estimation algorithms using gait phase information," *IEEE TBME*, vol. 59, no. 10, pp. 2884–2892, Oct 2012.

[45] J. Shotton et al., *Efficient Human Pose Estimation from Single Depth Images*. Springer, 2013, pp. 175–192.

[46] S. Bharathidason *et al.*, "Improving classification accuracy based on random forest model with uncorrelated high performing trees," *IJCA*, vol. 101, no. 13, pp. 26–30, Sep 2014.

[47] T. Weinkauf *et al.*, "Persistence1d: Extracting and filtering minima and maxima of 1d functions," https://www.csc.kth.se/~weinkauf/notes/persistence1d.html, Jan 2019.

[48] L. Breiman, *Classification and regression trees*. Routledge, 2017.

[49] VICON, "Plug-in-gait model," https://www.vicon.com/downloads/documentation/plug-in-gait-model-details, Jan 2019.

[50] T. Fawcett, "An introduction to roc analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861–874, 2006.

[51] L. L. Presti and M. L. Cascia, "3d skeleton-based human action classification: A survey," *Pattern Recognition*, vol. 53, pp. 130–147, 2016.